



*Advanced Dataflow Services for
Heterogeneous Multicore SoCs*

Drew Wingard

Sonics, Inc.

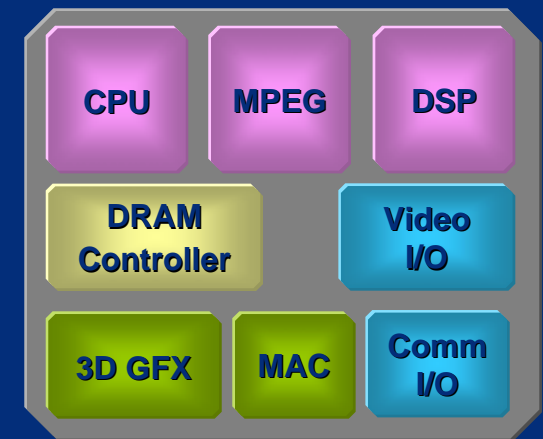
Mountain View, CA USA

wingard@sonicsinc.com

SoC Architecture Trends



- Massive feature integration
 - Driven largely by Moore's law (supply) and convergence (demand)
- Continued movement of complexity to software
- Distributed architectures
 - Higher scalability (and independence?)
- Multiple processors
 - CPU
 - DSP
 - Special purpose (MPEG, packet, ...)
- Distributed DMA
 - Removes centralized DMA bottleneck
 - Simplifies driver software integration



System On Chip

Multicore SoC Architecture Options

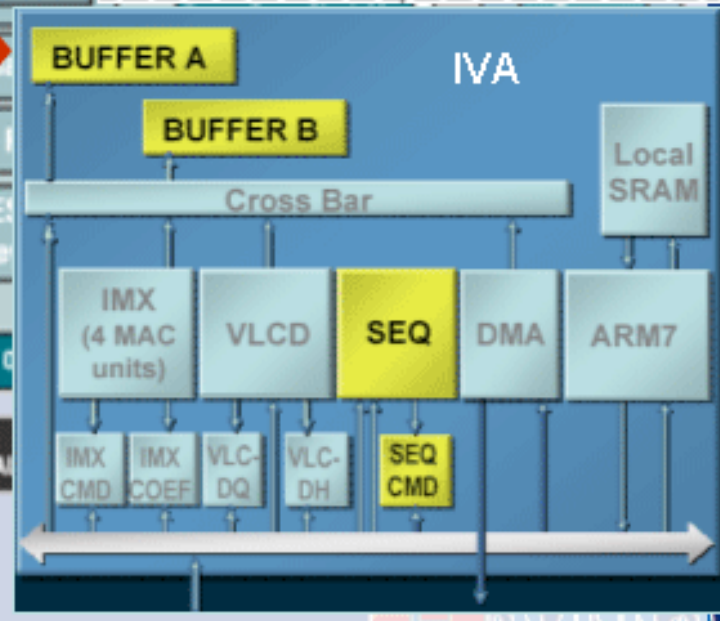
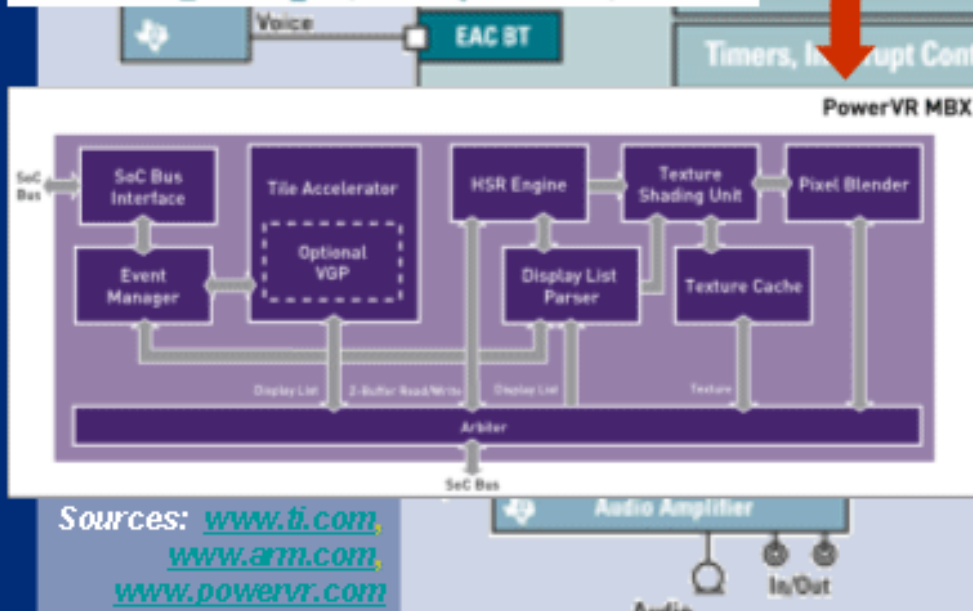
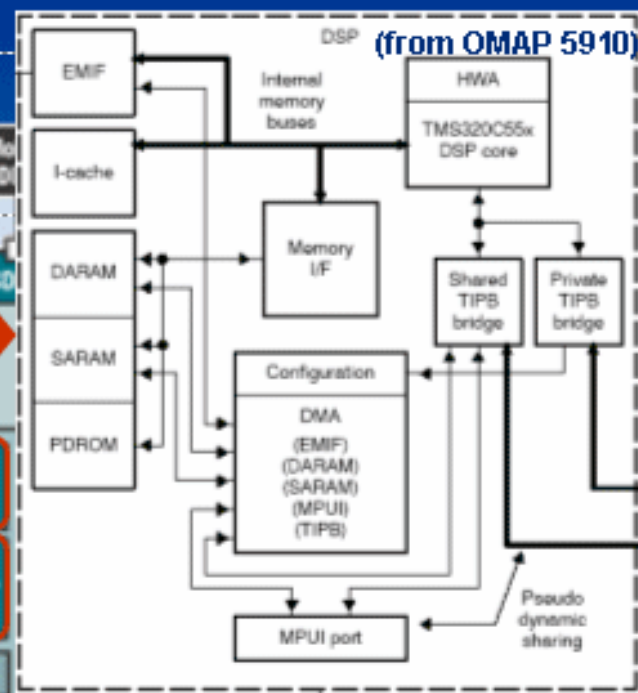
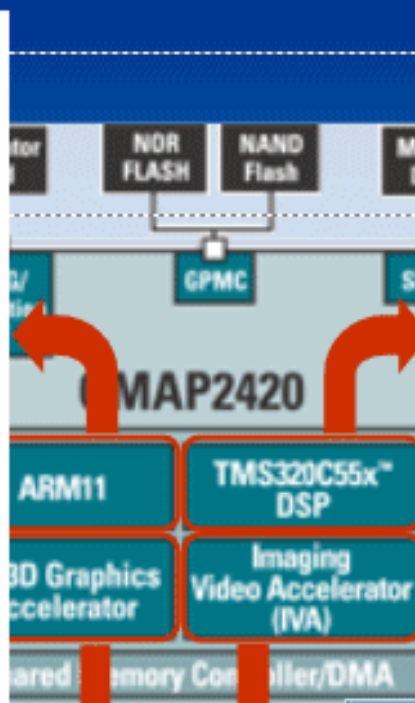
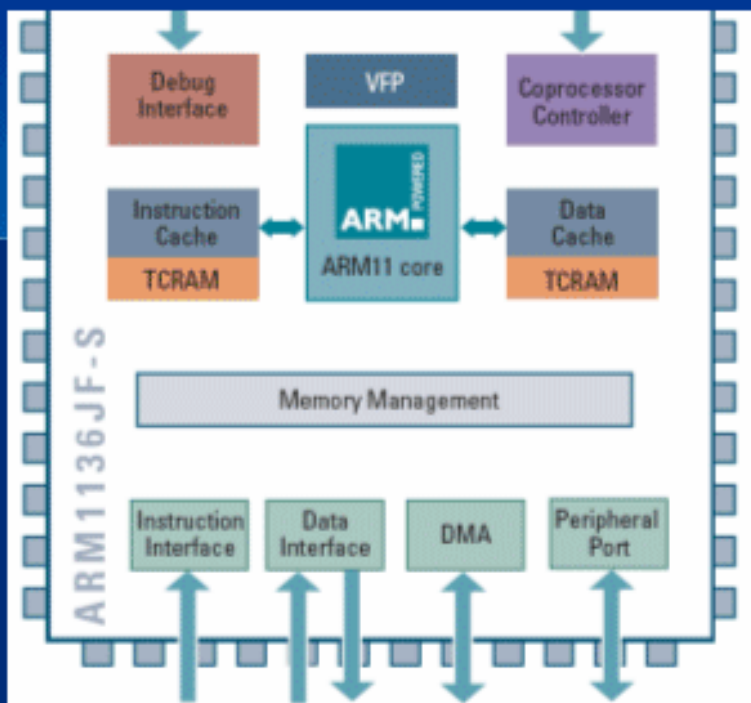


- Multiprocessor cluster
 - + Many operating systems know how to schedule, good for computing
 - Hard to scale past about 4 CPUs
 - Bad for real-time performance
- Uniform distributed computing fabric
 - + Highly scalable, locally efficient
 - Hard to program/schedule
 - Uniformity hurts Quality of Results (QoR)
- Distributed heterogeneous processing subsystems
 - + Per-subsystem mix of hard-wired and programmable
 - + Highest scalability, high QoR
 - + Divide-and-conquer programming model
 - + **Can deploy cluster and/or fabric approaches in subsystems**

Tile-Based Heterogeneous Multicore SoCs



- Tile – a distributed, largely independent subsystem for an SoC, normally composed of:
 - Processing
 - Memory
 - I/O
- Tile processing can be performed in fixed or programmable logic, or by general-purpose or special-purpose programmable processors
- Key elements of tile-based platforms:
 - Socket-based design
 - Decoupled interconnect architectures
 - Explicit capture of external memory/communication constraints
 - Delivery of required firmware together with tile hardware



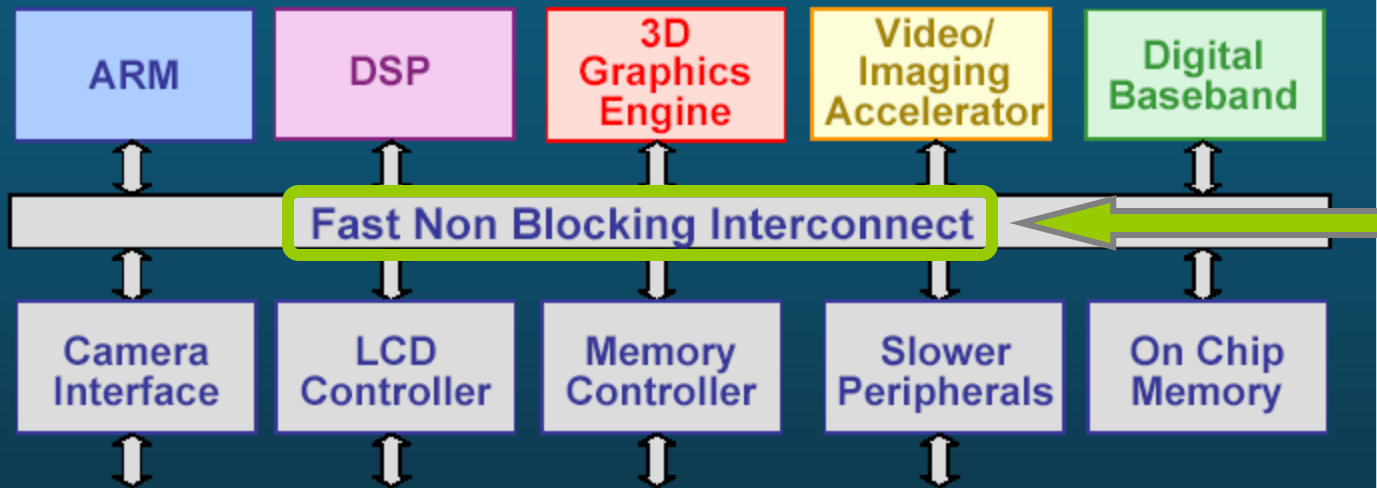
Sources: www.ti.com,
www.arm.com,
www.powervr.com

Multicore Architecture Advantages



Multi-Engine Processor Architecture for Parallel Processing

Multiple Engines Running Multiple Tasks



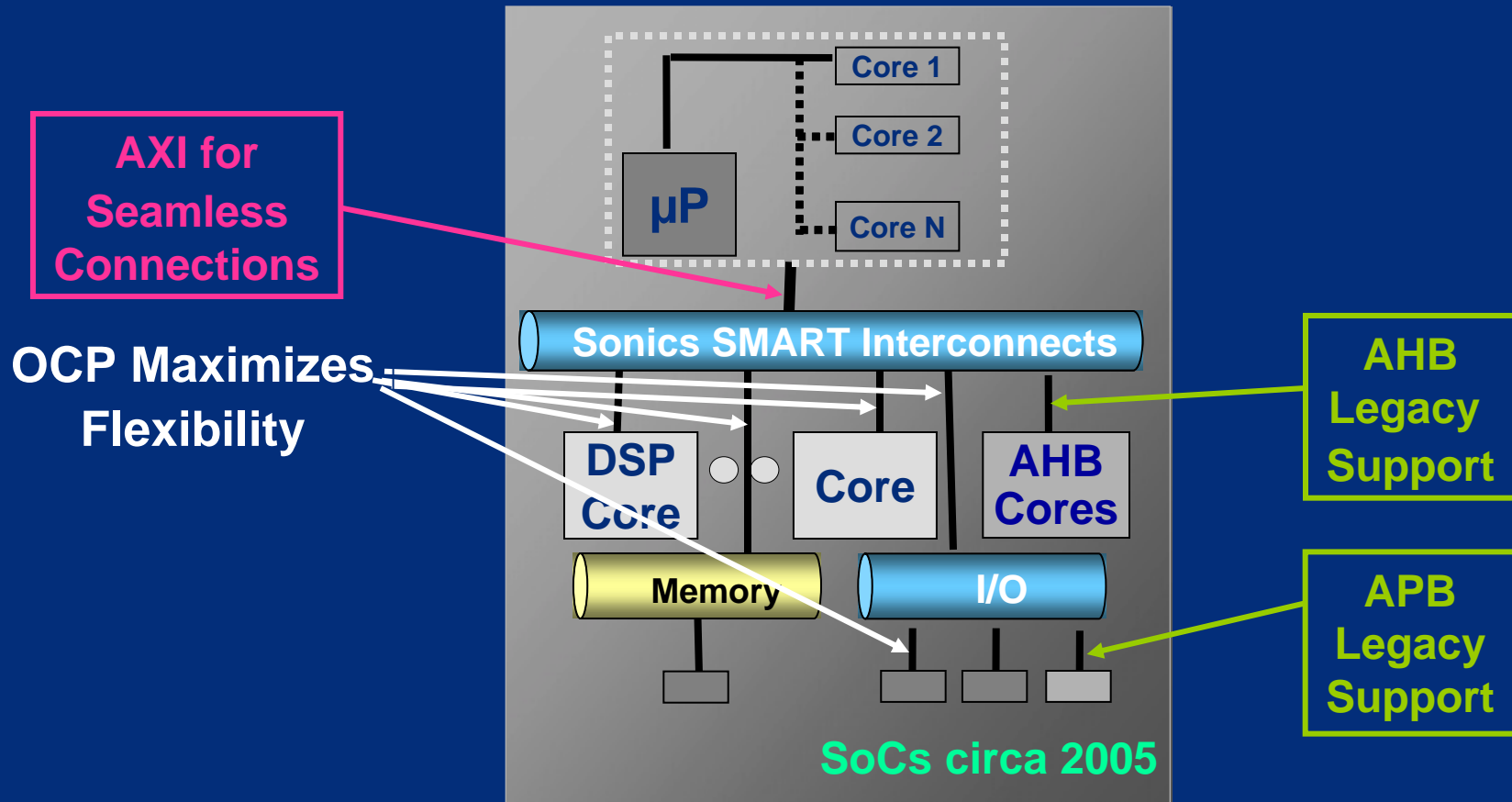
What is
needed

Simultaneous Parallel Processing

- ◆ No degradation in quality of service
- ◆ Highly responsive
- ◆ Flicker-free video and click-free audio during multitasking

Avner Goren
TI
EPF 2004

An Intelligent Interconnect Company



Sonics Adds Intelligent Dataflow Services



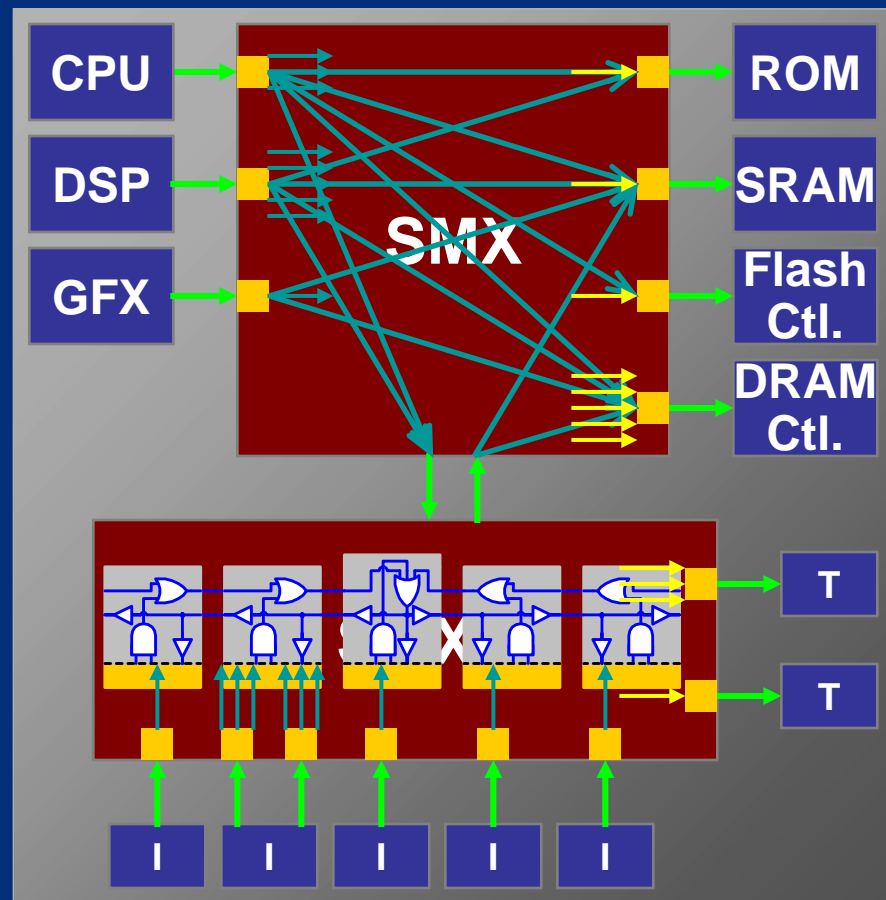


- Third-generation SMART Interconnect (network-on-chip, or NoC) from Sonics
- Targeted at mobile multicore SoCs
- Deployed by multiple customers in application processors
 - TI and Toshiba are announced users in mobile handset market
- Two key components
 - Advanced internal fabric
 - Intelligent agents

SonicsMX Basic Architecture



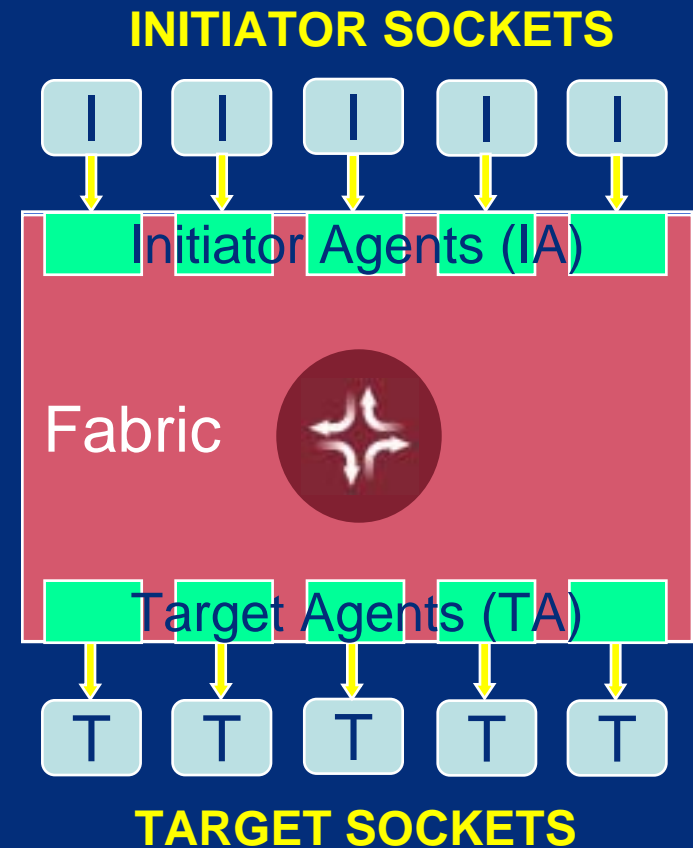
- Hybrid topologies
 - Full / partial cross bar
 - Shared bus
- Fully split (dual) request / response networks
- Pipelined, multithreaded, nonblocking fabric
- Distributed quality-of-service arbiter
 - Spans cycle, frequency, and data-width boundaries
 - Supports flexible thread-merging topologies



The Intelligence is in the Agents



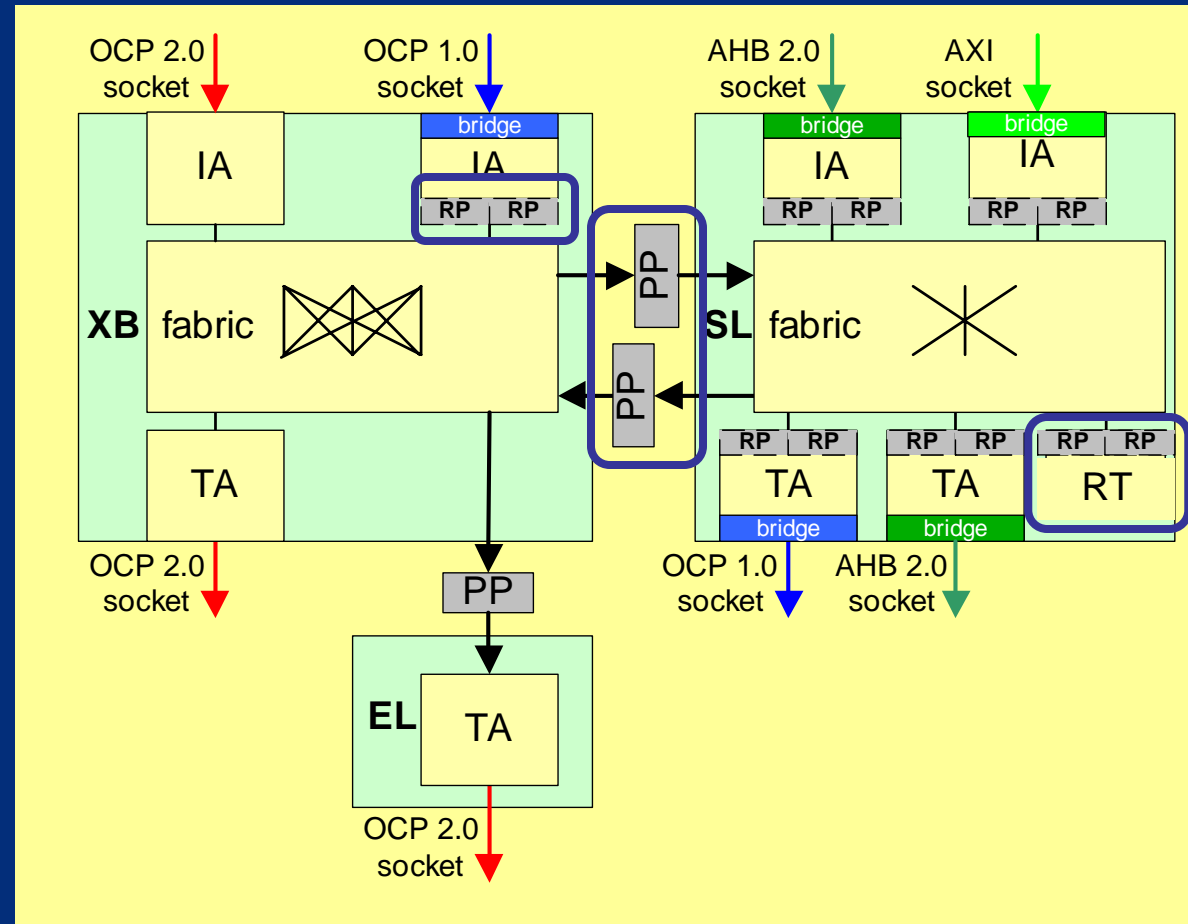
- Agents provide...
 - Protocol conversion
 - Agent adapts to IP core
 - Decoupling of IP cores from fabric
 - Provide local, isolated environment
 - Dataflow services
- Proven technology
 - Over 100 million ICs shipped so far
- Agent dataflow services
 - QoS-based arbitration
 - Power management
 - Access security
 - Error management
 - Burst, width, and command conversion



SMX Internal Structure



- Exchanges
 - Cross bar (XB)
 - Shared bus (SL)
 - Extender (EL)
- Pipelining options
 - Registering at socket interface
 - Register points (RP) at agent-fabric edge
 - Pipeline points (PP) between exchanges
- Register target (RT) to access SMX services
- Multiple socket support



QoS-Based Arbitration



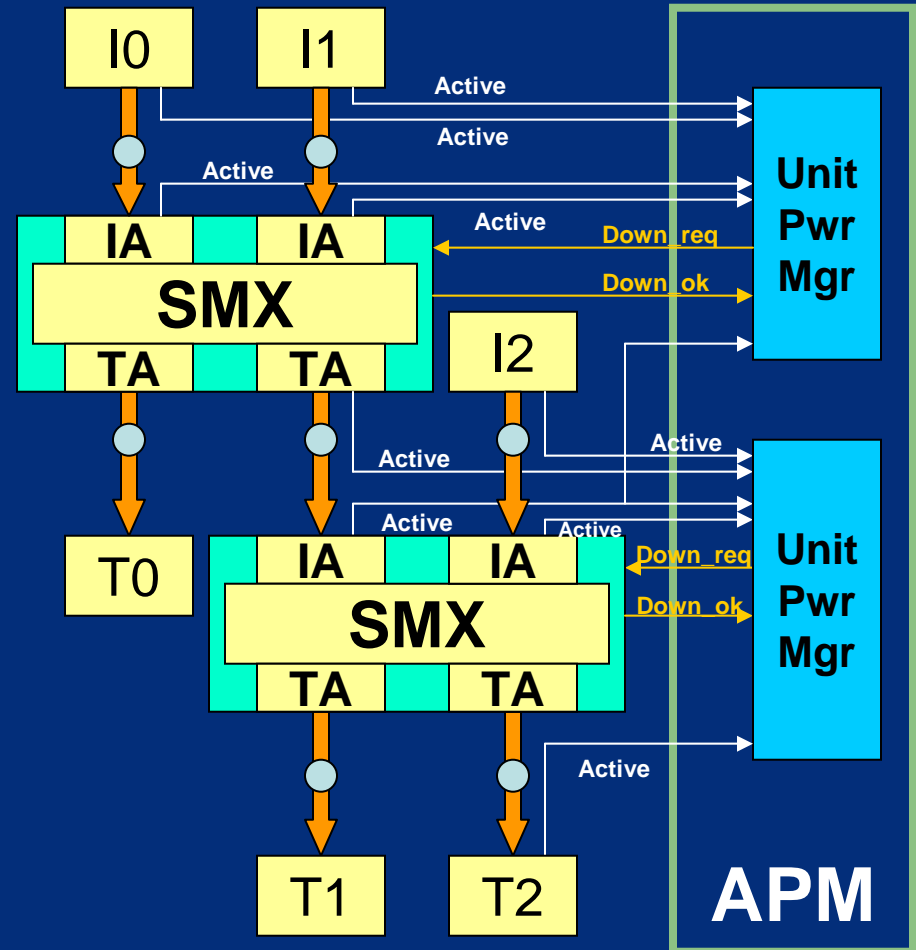
- Initiator dataflow threads mapped to target threads by SMX fabric
 - For example, 40 dataflows can share 8 DRAM threads in a digital video system
- Dataflows sharing a target thread are arbitrated using bandwidth weighting
- Independent threads are assigned to a QoS level (maintained throughout SMX)
- Nonblocking, multithreaded fabric and target interfaces allow:
 - Higher priority requests to interleave and respond before others
 - Guaranteed bandwidth threads to minimize buffering and receive latency guarantees
 - Optimum DRAM efficiency

Thread QoS Level	Bandwidth Allocation?	Quality of Service Model
Priority	Yes	Low latency while within bandwidth allocation, best-effort otherwise
Bandwidth	Yes	Guaranteed bandwidth while within bandwidth allocation, best-effort otherwise
Best-effort	No	(Not applicable)

Power Management System

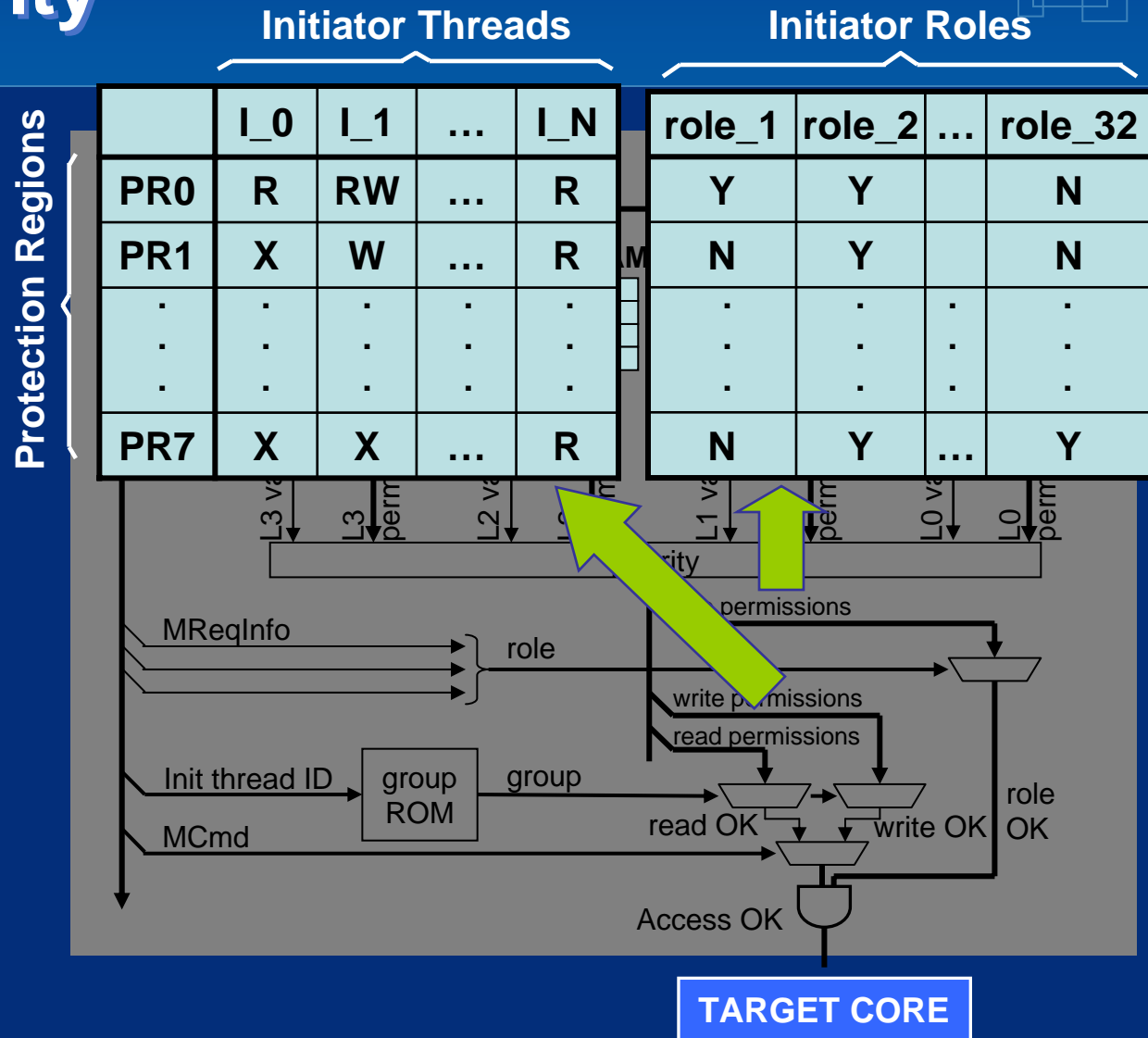


- Active status indication configurable on a per-socket basis
- Unit power manager (UPM) initiates power-down request when idle for some time
 - Based on system-defined policy
 - Can monitor subset of cores and SMX
- UPM initiates power-up request when attached master is active
 - Might be initiator or another SMX
- Supports chaining
- Simplifies design of application power manager (APM)



Access Security

- **Optional multiregion firewall**
 - Per-target, reprogrammable
 - Layered architecture supports rich set of security domains with variable region sizes
- **Access permissions determined per role and access type**
- **Flexible caching and reporting of security errors**

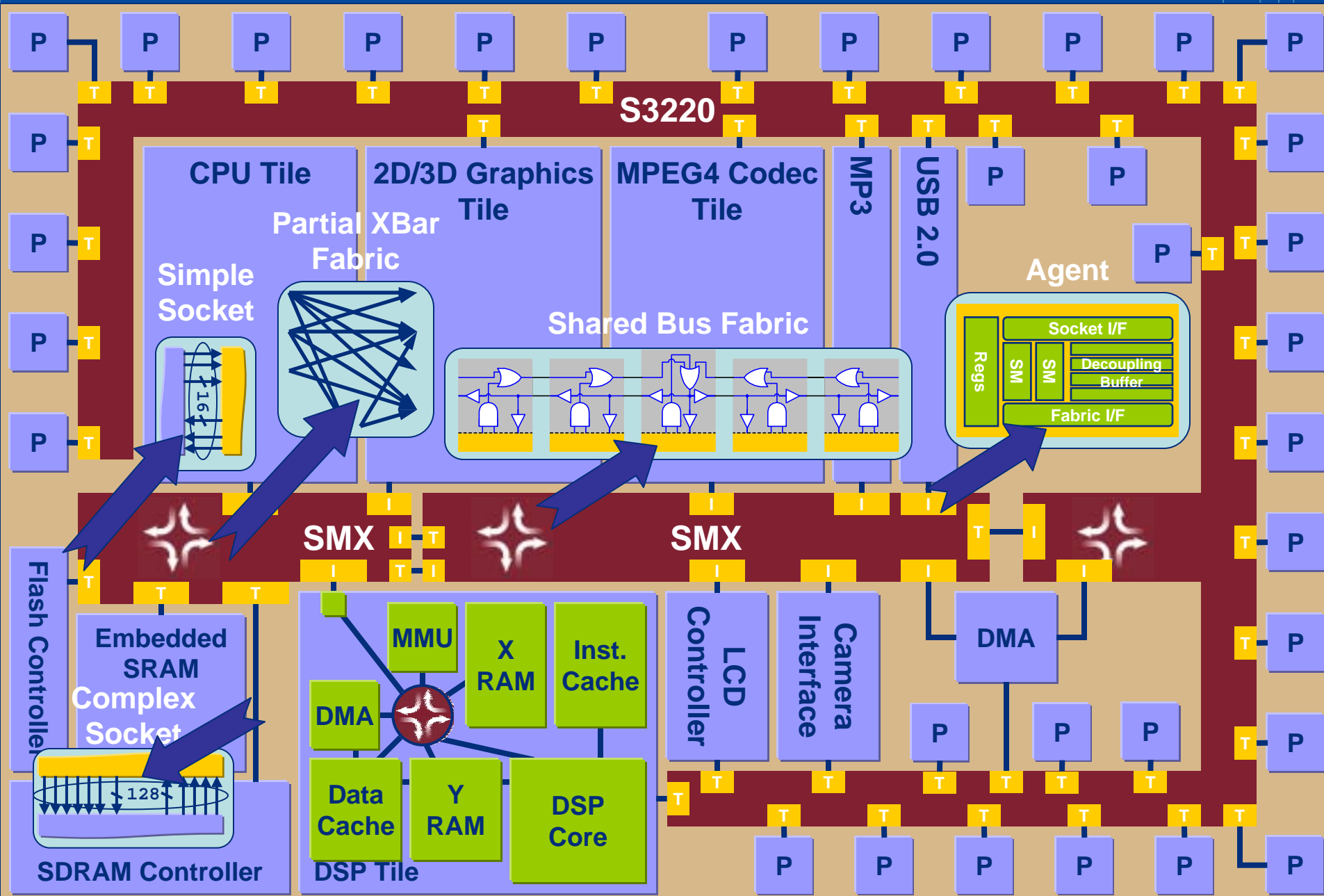


Error Management



- Detects a wide variety of error conditions
 - Bad addresses and illegal commands
 - Timeouts (initiator and target)
 - Security violations
- Aggregates errors (response and sideband) from IP cores
- Logs errors in agents for software interrogation
 - Status bits (i.e., have detected error, multiple errors)
 - Identifies initiator/address that caused error
- Reports errors as desired
 - In-band via error responses
 - Sideband via SError/MError, interrupts, etc.
- Supports IP-core isolation and reset for error recovery

Multicore Mobile Handset Example



Summary



- Multicore SoC designs are already common
 - And are largely heterogeneous
- Current approaches abstract each CPU into local tile
 - Increases independence of firmware for better scaling
- Intelligent interconnects are key to multicore SoCs
 - Nonblocking internal fabrics keep processors fed with data
 - Agent-based dataflow services are key to managing heterogeneity
- SonicsMX demonstrates the value of centralized dataflow services

Acknowledgements



- Entire SMX team, especially Joe Chou, Jeffrey Ebert, Scott Evans, Steve Hamilton, Michael Meyer, Ian Swarbrick, and Jay Tomlinson

Thank You